# A STUDY OF THE EFFECT OF NON-ABILITY VARIABLES ON THE OUTCOME OF INTERCOLLEGIATE DEBATES

By

Sidney Ray Hill, Jr.

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
1973

UNIVERSITY OF FLORIDA
3 1262 08552 9252

For Ruth

ACKNOWLEDGMENTS

The author wishes gratefully to acknowledge the assistance of the members of his supervisory committee. Professors Douglas G. Bock, Donald E. Williams, Anthony J. Clark, and David T. Hughes all gave freely of their time. Their encouragement and guidance were of major importance. The Department of Speech of the University of Florida provided numerous facilities as well as the financial support which made this dissertation possible.

A special note of thanks must go to two colleagues in the field of forensics: Mr. Glenn Pelham, former Director of the Barkley Forum at Emory University, and Mr. Brad Bishop, Director of Forensics at Samford University. Without their generous assistance in securing data, this study would have been much more difficult.

The greatest debt of all is owed to both his families, and most especially to his wife, Ruth. Without her interest, her encouragement, and her unfailing good humor, this dissertation would never have been completed.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS. . . . . iii

LIST OF TABLES. . . . . v

ABSTRACT. . . . . .vi

CHAPTER I: STATEMENT OF THE PROBLEM . . . . . 1

CHAPTER II: METHODOLOGY. . . . . .24

CHAPTER III: RESULTS. . . . . .43

CHAPTER IV: CONCLUSIONS. . . . . .65

BIBLIOGRAPHY. . . . . .84

BIOGRAPHICAL SKETCH . . . . . .89

LIST OF TABLES


Table I. Pearson Product-Moment Correlations for Dependent Measures of Outcome . . . . . 27

Table II. Tests for Significance of Pearson Product-Moment Correlations . . . . . 28

Table III. Regression Weights for Format Variables . . . . . 43

Table IV. Regression Weights for Format Variables On Ballot Measures . . . . . 45

Table V. Percentages of Excessive Negative Wins . . . . . 46

Table VI. Regression Weights for Rater Variable Sex . . . . . 47

Table VII. Regression By Sex on Win-Loss for Groups . . . . . 48

Table VIII. Regression of Sex on Win-Loss . . . . . 49

Table IX. Regression of Sex on Speaker Ratings . . . . . 50

Table X. Regression of Sex on Team Ratings for Groups . . . . . 51

Table XI. Regression of Sex on Team Ratings . . . . 52

Table XII. Regression of Sex on Speaker Rankings for Groups . . . . . 52

Table XIII. Regression of Sex on Speaker Rankings . . 53

Table XIV. Regression Weights for Rater Variable Prestige . . . . . 55

Table XV. Regression of Prestige on Win-Loss . . . 56

List of Tables, continued . . .


| | | |
|---|---|---|
| Table XVI. | Regression of Prestige on Speaker Rating . . . . . | 57 |
| Table XVII. | Regression of Prestige on Team Ratings For Groups . . . . . | 58 |
| Table XVIII. | Regression of Prestige on Team Ratings . . . . . | 60 |
| Table XIX. | Regression of Prestige on Speaker Rankings . . . . . | 61 |
| Table XX. | Regression Weights for Rater Variable Proximity . . . . . | 62 |
| Table XXI. | Regression of Proximity on Speaker Ratings . . . . . | 63 |
| Table XXII. | Regression of Proximity on Team Rating . . . . . | 64 |

Abstract of Dissertation Presented to the Graduate Council of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

# A STUDY OF THE EFFECT OF NON-ABILITY VARIABLES ON THE OUTCOME OF INTERCOLLEGIATE DEBATES

by

Sidney Ray Hill, Jr.

August, 1973

Chairman: Douglas G. Bock
Major Department: Speech

This study examines the effect of two format variables, side of topic and speaker position, and three rater variables, sex, prestige, and proximity, on the outcome of intercollegiate debates. The data were gathered from three different debate tournaments covering a five year period. In all, the data pool consisted of two thousand, five hundred and seventy different debates. Using Mendenhall's procedure for achieving limits on the errors of estimation, samples were drawn from the data pool and submitted to analysis by the University of Florida Computing Center, using the BMDX63 "Multivariate Linear Hypothesis" program.

The results of the regression analysis indicate that the format variables had no significant effect on the outcome of debates studied.

However, in the case of the rater variables studied, significant regression effects were found. In general, male debaters had a higher expectation of winning any given debate, and received higher speaker ratings than did female debaters. Interaction effects between the sex of the debater, the sex of the debater's colleague, and the sex of the judge were observed.

Studies of the rater variable prestige indicated that high prestige teams receive higher scores on speaker rating, team rating, and speaker ranking than do either low prestige or junior college teams. Again, interaction effects between the prestige of the debaters, the prestige of their opponents, and the prestige of the judge were observed.

Studies of the rater variable proximity revealed that teams farthest from the judge geographically have a higher expectation of winning and receive higher scores than their opponents.

The results of this study allowed the development of a number of prediction equations whereby outcome on any of the A.F.A. Form "C" ballot measures may be predicted on the basis of the rater variables.

# CHAPTER I

## STATEMENT OF THE PROBLEM

Intercollegiate debate operates under a format of rigidly limited, alternating speeches by opposing two-man teams. The teams uphold opposite sides of a specified resolution. A critic judge renders a decision in favor of the team which has done the better debating. The process is not intended to discover "truth" in the resolution. Rather it is designed as a neutral framework within which the debaters may display their argumentative skills.

The purpose of this study was to test the neutrality of the debate format, and to determine the influence of certain non-ability variables on the decisions rendered by debate judges. The procedure for the study was a multi-variate regression analysis of data obtained from A.F.A. Form "C" debate ballots.

There seems to be ample basis for the statement that intercollegiate forensics faces widespread criticism from various positions within the field of speech. Certainly the proceedings at the Chicago convention of the Speech Communication Association and the American Forensics Association support this observation.[1] In the

face of such attacks, members of the forensics profession have responded vigorously. Numerous articles in scholarly journals as well as convention programs and addresses have defended the legitimacy of forensics within the discipline of speech and its value in higher education.[2] A survey of those efforts at defense, however, reveals that they all assume the validity of the current practices with forensics. Only occasionally does research focus on specific problems within intercollegiate forensics, and when such focus does occur, the studies are limited in scope and in application.

Intercollegiate forensics in the 1979's is a major academic activity. Vast numbers of students and faculty personnel are involved. Large amounts of money are expended to travel hundreds of thousands of miles to participate in tournaments. In 1972-73 there were over 300 different tournaments held. In view of this, one would assume that the basic format under which all of this activity was held would have been the object of extensive study. It seems reasonable to expect, but in fact that is not the case. Research in forensics has tended for the most part to concentrate on demographic surveys[3] and bitter squabbles over technical points of argumentation.[4] While not denying the worthiness of such studies, it should be pointed out that they do not deal with any question of

validity of the process itself. They assume that the process is working to achieve the desired results. Whether warranted or not, such an assumption needs to be tested.

It seems appropriate at this point to make some mention of the vast body of published and graduate research annually reported in the American Forensics Association bibliographies. The bulk of such work consists, in fact, of historico-critical studies of public and political debates and debaters. The next largest catagory consists of essays on the theory of argumentation. Only a small percentage of the research reported each year is of a quantitative nature, and only a minor portion of that deals with intercollegiate forensics.[5]

There is evidence to suggest that this breakdown adequately describes the entire body of forensics research. An analysis and summary of graduate research in debate through 1950[6] reported eighty-six various theses and dissertations. 38% of these were historico-critical studies or political debates and debaters. An additional 18% consisted of histories of debating and of debate societies at various institutions. 29% were descriptive studies of debate, debate texts, and debaters. Two studies dealt with the philosophy of argumentation, and five dealt with the personality traits of debaters -- primarily with critical thinking ability. Only one

graduate study prior to 1950 attempted to measure, even indirectly, the balance of the traditional debate format. Cromwell in 1949 examined order effect and discovered no difference in audience effectiveness attributable to the order in which two opposing speeches were presented.[7]

The most sizable group of studies which attempt to identify and analyze variables effecting the outcome of intercollegiate debates focus on ability variables. These may be defined as those factors commonly recognized to be legitimate bases for the awarding of decisions in competitive debates. They include analysis, argument, reasoning, case and organization, evidence, refutation, and delivery; the exact list being dependent on the actual ballot in use.

Past studies have utilized survey sampling to determine which of these criteria were weighed most heavily by judges in making their decisions.[8] The use of evidence in debate has long been an object of study. Several researchers have considered the role of evidence in the theory of argumentation.[9] The use of evidence in intercollegiate debate has been examined by a number of researchers. Studies dealt with such factors as frequency of use, types of evidence used, and accuracy in the citation of evidence.[10]

The relationship, if any, between debate and critical thinking ability has received wide study, with various

conflicting results. Phillips, in 1962, could find no evidence that participation in debate actually improved critical thinking ability.[11] He explained differences between his findings and those of Howell and of Brembeck[12] as caused by the fact that intercollegiate debaters tend initially to rank above the average in critical thinking ability. Phillips' findings were disputed by Faules, whose 1967 study argued that skill in refutation and critical thinking ability were manifestations of the same trait.[13]

The influence of refutation skills on success in debate was the subject of study for Keeling in 1959.[14] He discovered evidence to suggest that refutation skill might well be sufficiently significant to serve as a predictor for debate effectiveness.

This finding was generally supported by the results obtained by Wise in 1971.[15] Through examination of debate ballots, Wise discovered various relationships between scores on selected ballot traits and win-lose results for a given debate round. Specifically, the category of "analysis" for negative speakers and that of "refutation" for affirmative speakers was found to be crucial in determining win-lose.

An interesting extension of his examination by Wise revealed, first, that affirmative speakers tend to score higher than their negative opponents, even though the

win-lose record is balanced between affirmative and negative.[16] Wise explained this difference on the basis of consistently higher affirmative scores on the ballot categories of "organization" and "delivery." In his sample, this held true for both winning affirmatives compared to winning negatives as well as losing affirmatives compared to losing negatives.

These data suggested the need for further study to determine whether or not the present rating system does in fact provide a balanced opportunity for negative as well as affirmative speakers. If the trend reported by Wise should be substantiated, it would clearly establish the necessity for some revision of the present system where each speaker, affirmative and negative, is scored on a 5-30 point scale.

In addition to those descriptive studies which set out to categorize debaters in terms of certain personality traits, there have been a small number of studies which have attempted to test statistically the relationship between certain traits and success in debate. Two in particular are worth mentioning. Allen, in 1963, studied the variables of interpersonal and concept compatibility as they related to the success of a debate team.[17] He reported that interpersonal compatibility between members of a debate team resulted in more efficient as well as more effective case construction, while concept

compatibility resulted only in greater efficiency. These data did not reveal that student satisfaction with the debate experience was any greater as a result of pairing with a compatible colleague.

Willmington, in 1969, tested the relationships between knowledge of debate theory, amount of previous debate experience, and debate effectiveness.[18] His results indicated that the only significant predictor of success would be amount of previous debate experience.

While each of the studies cited above provides worthwhile data, it should be clear at this point that none of them directly approached the study of the internal validity of the debate process. Only Wise provides any evidence as to the adequacy of balance in the currently used 10-5 format, and this was only a minor and indirect result of Wise's study. No evidence has yet been presented on the issue of judge reliability, or even of judging bias. One might infer from this negative evidence that the persons who are involved in debate are relatively well satisfied with the way in which the activity is currently structured. This is simply not the case. A number of studies suggest that coaches as well as student debaters frequently question the accuracy of decisions rendered by the judge of a particular round. Kruger and Hufford have examined the question of the consistency of decisions across judges.[19] Two researchers reported results which

might lead to methods for checking the accuracy with which judges determine win-lose decisions.[20] Other studies have focused on the qualifications of judges as well as the methods by which they arrived at decisions.[21]

All of these studies support the assertion that doubts exist concerning debate judging. Recently two studies have attempted to resolve those doubts. Brooks, in 1971, focused on the phenomenon of geographical distance bias in debate decisions.[22] As Brooks explained, the proximity of the judge's school to their opponent's school is an excuse frequently offered by losing debaters.[23] The study reported plotted the results of six mid-western debate tournaments in terms of wins by nearest or farthest teams. At least a 10% difference in mileage was established as the minimum definition of difference in distance, and those teams not meeting this definition were defined as equidistant from the judge's school and excluded from this study. Using a Chi-Square statistic with a .05 level of confidence, Brooks found that in 5 of the 6 tournaments studied, the team geographically nearest to the judge's school won the most debates.[24] Brooks' results are extremely significant in that they suggest that tournament evaluation may not be strongly related to the stated criteria under which decisions are supposed to be made.

However, this study does not extend far enough. First, it fails to consider the possibility of effects due to social distance. In intercollegiate debate, social distance would manifest itself as a prestige factor generated by the quality of the debating performed, and the academic and forensics reputations of the schools represented. Also involved would be the experience and reputation of the individual students and judge present.

The division of the country into various districts for purposes of selection to the National Debate Tournament introduces another type of "distance" which Brooks does not consider. Schools normally do a major portion of their season's debating within their N.D.T. district, thus potentially fostering "friendship through propinquity" as Brooks suggests. However, competition for N.D.T. invitations is quite intense, and rivalries are established which might well over-ride such friendship. All of these points suggest the need for further examination of the variable of distance as it affects the outcome of intercollegiate debates.

The second study of interest was reported by Hayes and McAdoo in 1972.[25] Stimulated by the work of Hensley and Strother,[26] Hayes and McAdoo set out to examine the variable of sex as it affected evaluation in intercollegiate debate. The study utilized data generated by speaker rankings rather than the win-lose results employed by

Hensley and Strother in an attempt to isolate individual rather than team evaluations. A Chi-Square test found significant deviation from expected results at a .01 level of confidence, and the direction of the results indicated that female debaters in the sample were evaluated more highly than were male debaters.

There are a number of limiting factors in the work reported by Hayes and McAdoo. First was the method by which they collected their sample. The ballots studied included all of those accumulated by three different college debate programs over a three-year span. Obviously this resulted in multiple measures of the same female and male debaters. While this procedure may give conclusive evidence of the superiority of female debaters at the schools involved, there is no basis for generalization to the entire population of female debaters. Second, the experimenters excluded from the data pool all ballots in which the competition consisted of all men or all women. Thus the question of interaction between male and female debaters is raised. Finally, the study made no distinctions as to the sex of the judge, thus allowing forthe possibility of yet another contaminating variable; i.e. interaction between the sex of the judge and sex of the debater evaluated.

On the basis of the studies examined above, it is suggested that there is a need for systematic examination

into non-ability variables which may affect the outcome of intercollegiate debates and the evaluations received by the students participating in those debates.

An examination of the A.F.A. Form "C" ballot indicates that included among the data provided by this instrument is the following descriptive and evaluative information:

Descriptive:

1) The name of the judge
2) The current school affiliation of the judge
3) The schools represented in the round
4) The names of the individual debaters

Evaluative:

1) The side (and school) which won the round
2) Ratings for each of the four debaters on a 1-5 scale on 6 traits
3) Rankings of the four debaters from 1st to 4th
4) Ratings for each team on a scale of 1-5 points

In addition to the ability variables measured by the evaluative data on the ballot, it is possible to isolate the following non-ability variables.

Variable #1 - Side of the topic

The format of intercollegiate debate is deliberately designed to eliminate any advantage accruing to either the affirmative or negative side within the debate.[27] While the affirmative carries the burden of proof, it is given the right to select the specific areas in which to argue the debate. The negative enjoys the presumption, but is required to respond directly to the affirmative analysis.

The affirmative is given the first and last speeches, but the negative has a fifteen-minute block in the middle of the debate. Actual speaking time is balanced between the two sides. Normally, teams "switch sides" from affirmative to negative and back in the course of a tournament, but an equal number of affirmative and negative rounds is not required. Tournament team awards are based on total win-loss records, and individual awards on cumulative speaker points, with no consideration as to the side of the topic on which the team or student debated. All of these practices depend for their validity on the assumption that there is no inherent advantage in either the affirmative or negative side. Yet systematic research on the point is not available. Some test of the assumption of balance between sides seems, therefore, to be desirable. In order to examine the assumption, the following research hypotheses will be utilized:

1.1 There is no effect due to side in the win-loss decisions of intercollegiate debate.

1.2 There is no effect due to side in the individual speaker ratings in intercollegiate debates.

1.3 There is no effect due to side in the team ratings in intercollegiate debates.

Variable #2 - Speaker Position

While the theory of competitive debate assigns to each speaker within the debate specific duties and

responsibilities, there is no consensus that any of these duties is significantly more difficult than any other. While the first affirmative constructive speech is prepared in advance, thus freeing that debater from the demands of extemporaneous adaptation, the first affirmative rebuttal speech is generally considered to be the most difficult of the eight speeches. Again, the format seeks balance between the speakers. On the assumption that this balance is achieved, all four speakers are rated on the same 5-30 point scale, and are compared to each other in a forced choice ranking from 1st to 4th. Moreover, individual speaker awards at tournaments are made from the entire pool of participants, with no consideration given to the order in which the students spoke. Previously cited research by Wise raised questions against the assumed balance between speaker positions. In order to validate that assumption, the following research hypotheses will be tested:

2.1 There is no effect due to speaker position in the rating scores of affirmative debaters.

2.2 There is no effect due to speaker position in the rating scores of negative debaters.

2.3 There is no effect due to speaker position in the rankings of affirmative debaters.

2.4 There is no effect due to speaker position in the rankings of negative debaters.

Variable #3 - Sex

With the growth of co-education in institutions of higher learning, women have participated more and more in debate against men. The tournament with separate men's and women's divisions is rapidly disappearing. Women have participated in the National Debate Tournament and have compiled outstanding records. Some of the outstanding coaches in the country are women. Women have served as president of the major forensics honoraries, and play an active part in the professional societies. With this background, one would have no reason to believe that sex might serve as a variable affecting the outcome of intercollegiate debates. Yet many coaches do hold such a belief. Hayes and McAdoo reported that "Many debate coaches throughout the country hold the belief that male debaters are superior to the female debater."[28] Hensley and Strother reported results indicating that sex does affect win-loss decisions, with a mixed (i.e. one male, one female) team standing a "greater than random chance of winning."[29] In light of these facts, further research seems necessary. Accordingly, the following research hypotheses will be tested:

3.1 There is no effect due to sex in win-loss decisions in intercollegiate debates.

3.2 There is no effect due to sex in speaker ratings in intercollegiate debates.

3.3. There is no effect due to sex in speaker rankings in intercollegiate debate.

3.4 There is no effect due to the sex of the judge in win-loss decisions in intercollegiate debates.

3.5 There is no effect due to the sex of the judge in speaker ratings assigned in intercollegiate debates.

3.6 There is no effect due to the sex of the judge in speaker rankings assigned in intercollegiate debates.

3.7 There is no effect due to an interaction between the sex of the judge and the sex of the debaters in the outcome of intercollegiate debates.

Variable #4 - Prestige

The subjective nature of decision-making in intercollegiate debate has frequently led to questions concerning the impact of prestige in determining those decisions. The tournament circuit generally results in frequent contacts between teams, and a widespread

knowledge of the past successes of any team or teams. Do judges favor certain teams over others? Or, perhaps, teams from certain schools? The problem is compounded by the fact that prestige tends to be generated in forensics circles by success, and is not necessarily related to the general academic reputation of an institution nor to the grade point average of the individual debater.[30] How much of a debater's success is attributable to his ability, and how much to the prior reputation with the judge(s)? How much of the success which teams from a given school achieve year after year is due to an outstanding forensics program turning out superior teams? And how much must be credited to the "mantle of greatness" which any one team enjoys because of the achievements of past teams? The fact that these are difficult questions to answer does not mean that they should be avoided.

The first step in searching for those answers requires that some definition of prestigeous teams be formulated. For purposes of this study, "high prestige" teams will be defined as those representing an institution which attended the National Debate Tournament either in the current year or the year immediately preceding the one in which the debate studies was held. "Low prestige" teams will include all other teams representing four-year institutions. Junior Colleges will be considered in a separate category

because of the widespread assumption that debaters from junior colleges have less ability than those from senior institutions.

In order to equalize ability of the debaters included in the sample insofar as possible, the hypotheses testing variable #4 will be tested using only data from elimination rounds. Because teams selected to participate in elimination rounds must qualify in multi-round preliminaries, and because of the practice of power-matching teams in those preliminary rounds, limiting the data pool to elimination round ballots will remove from consideration clearly inferior teams. The over-all variance in team and individual ability should be greatly reduced by such a procedure, thus increasing the reliability of whatever test results are obtained. The research hypotheses to be employed are:

4.1 There is no effect due to prestige in win-loss results of debates between senior colleges.

4.2 There is no effect due to prestige in the win-loss results of debates between "high prestige" colleges and junior colleges.

4.3 There is no effect due to prestige in the win-loss results of debates between "low prestige" colleges and junior colleges.

Variable #5 - Proximity

As was discussed above, geographical proximity has frequently been suggested as a variable influencing the outcome of intercollegiate debates. Brooks offers evidence to suggest that judges vote more often in favor of teams from schools nearest to their own than they do for teams from schools further away.[31] Such a trend, if confirmed, represents a serious failing in the present judging system, and therefore deserves further examination.

An additional measure of proximity might be provided by considering the nine N.D.T. districts into which the nation is divided. Because these district lines tend to represent natural lines of travel and traditional rivalries, the effects due to simple geographical proximity might well be over-ridden by the pressures of district loyality. A judge from Mississippi, for example, belongs to N.D.T. District VI, which includes Alabama, Tennessee, Kentucky, Georgia, Florida, and the Carolinas. Such a judge might well feel more affiliation with, or be swayed more by friendship for, a team from North Carolina than for a team from Arkansas or Louisianna, in spite of the fact that he is geographically closer to those teams. The strength of district loyalties is demonstrated by the fact that the National Debate Tournament does not allow a judge to judge any team from his own district in competition against a team from another district. The

testing of variable #5, then, should allow for the consideration both of geographical and district proximity. The following research hypotheses have therefore been framed:

5.1 In debates where both teams and the judge represent the same N.D.T. district, there is no effect due to geographical proximity in determining the result.

5.2 In debates where both teams represent one N.D.T. district and the judge represents another district, there is no effect due to geographical proximity in determining the result.

5.3 In debates where each team and the judge represent different N.D.T. districts, there is no effect due to geographical proximity in determining the result.

5.4 In debates where one team and the judge represent the same N.D.T. district, and the second team represents another district, there is no effect due to district proximity in determining the result.

## NOTES - CHAPTER I

[1] See "Minutes of the 24th Annual American Forensics Association Meeting," Journal of the American Forensics Association, IX (Winter, 1973), 380-391.

[2] See, for example, Joseph W. Wenzel, "Campus and Community Programs in Forensics," Journal of the American Forensics Association, VII (Spring, 1971), 253-259; Donald G. Douglas, "Toward a Philosophy of Forensic Education," Journal of the American Forensics Association, VIII (Summer, 1971), 36-41; and the Spring, 1972 issue of the Journal of the American Forensics Association, containing articles on the same theme by Douglas, Henry McGuckin, Lee Granell, Steven Shiffin, and Robert D. Kelly.

[3] Such as Wayne E. Hensley, "A Profile of the N.F.L. High School Forensic Director," Journal of the American Forensics Association, IX (Summer, 1972), 282-287; or Allan J. Kennedy, "Directory of Universities and Colleges Conducting Summer Institutes," Journal of the American Forensics Association, VII (Winter, 1971), 224-233.

[4] See, for example, David A. Ling and Robert V. Seltzer, "The Role of Attitudinal Inherency in Contemporary Debate," Journal of the American Forensics Association, VII (Spring, 1971), 271-277; John F. Cragen and Donald C. Shields, "The Comparative Advantage Negative," Journal of the American Forensics Association, VII (Spring, 1970), 85-91; or Arthur N. Kruger, "The Comparative Advantage Case: A Disadvantage," Journal of the American Forensics Association, III (September, 1966), 204-211.

[5] The latest edition was published in the Journal of the American Forensics Association, VIII (Fall, 1971), 81-104. Not counting cross-references, it contains 288 entries. 120 of these were for studies of public and political debates and debaters. There were 53 essays in the theory of argument, persuasion, logic, and reasoning. Ten entries were for histories of debate and debate societies; 14 for essays on coaching; and 19 for articles on tournament and contest administration. There were 19 reported studies in attitude change and seven on decision making. There were 17 entries from law journals on rules of legal evidence, jury decision-making, etc.

[6] Beatrice Bahr Mills, "A Survey of Graduate Research in Debate" (Unpublished M.A. Thesis, University of Southern California, Los Angeles, California, 1952).

[7] Harvey Cromwell, "The Relative Effect on Audience Attitude on the First Versus the Second Argumentative Speech of a Series" (Unpublished Ph.D. Dissertation, Purdue University, Lafayette, Indiana, 1949).

[8] See David L. Dollar, "An Examination of the Importance Assigned Criteria of Debate Evaluations" (Unpublished M.S. Thesis, Kansas State Teachers College, Emporia Kansas, 1968); and Sandra Madsen, "A Study of the Use of Specified Criteria . . . in Evaluating Tournament Debates" (Unpublished M.A. Thesis, Wisconsin State University, 1969).

[9] Included among these are: Harold E. Smith, "The Use of Statistical Evidence in Debate," Quarterly Journal of Speech, XXVI (October, 1940), 426-431; Frederick George Marcham, "Teaching Critical Thinking and the Use of Evidence," Quarterly Journal of Speech, XXXI (October, 1945), 362-368; and Robert S. Cathcart, "An Experimental Study of the Relative Effectiveness of Four Methods of Presenting Evidence," Speech Monographs, XXII (August, 1955), 227-223.

[10] For examples of such studies see: William R. Dresser, "The Use of Evidence in Ten Championship Debates," Journal of the American Forensics Association, I (September, 1964), 101-106; Carl E. Larson and Kim Giffin, "Ethical Considerations in the Attitudes and Practices of College Debaters," Journal of the American Forensics Association, I (September, 1964), 86-90; Robert P. Newman and Keith R. Sanders, "A Study in the Integrity of Evidence," Journal of the American Forensics Association, II (January, 1965), 7-13; and James A. Benson, "The Use of Evidence in Intercollegiate Debate," Journal of the American Forensics Association, VII (Spring, 1971), 260-270.

[11] Gerald M. Phillips, "Experimentation and the Future of Debate," The Gavel, XLV (November, 1962), 3-6.

[12] William Smiley Howell, "The Effects of High School Debating on Critical Thinking," (Unpublished Ph.D. Dissertation, University of Wisconsin, Madison, Wisconsin, 1942); and Winston L. Brembeck, "The Effects of a Course in Argumentation on Critical Thinking Ability" (Unpublished Ph.D. Dissertation, University of Wisconsin, Madison, Wisconsin, 1947).

[13] Don Faules, "Measuring Refutation Skills: An Exploratory Study," Journal of the American Forensics Association, IV (Spring, 1967), 47-52.

[14] Russell M. Keeling, "An Analysis of Refutation and Rebutal in Interscholastic Debate " (Unpublished M.A. Thesis, Baylor University, Waco, Texas, 1959).

[15] Charles N. Wise, "Relationships Among Certain Rating Judgments on the Form C Ballot," Journal of the American Forensics Association, VIII (Spring, 1971), 305-308.

[16] Wise, p. 306.

[17] R.R. Allen, "The Effects of Interpersonal and Concept Compatibility on the Encoding Behavior and Achievement of Debate Teams," Central States Speech Journal, XIV (February, 1963), 23-26.

[18] S. Clay Willmington, "A Study of the Relationships of Selected Factors to Debate Effectiveness," Central States Speech Journal, XX (Spring, 1969), 36-39.

[19] Arthur N. Kruger, "Judging the Judging at Meadville," Bulletin of the Debating Association of Pennsylvania Colleges, XXI (December, 1955), 32; and Roger Hufford, "Toward Improved Tournament Judging," Journal of the American Forensics Association, II (September, 1965), 120-125.

[20] See Robert Shrum, "Do Judges Know When Debaters Win or Lose?" Rostrum, XLII (1968), 6-7; and Sidney R. Hill, Jr., "A Study of Participant Evaluations in Debate," Journal of the American Forensics Association, IX (Winter, 1973), 371-377.

[21] For examples of these, see: Dorothy Melzer, "Suggestions for Improving Debate Judging," Southern Speech Journal, XVIII (1952), 51-63; and Dean S. Ellis and Robert Minter, "How Good Are Debate Judges?" Journal of the American Forensics Association, IV (Spring, 1967), 53-56.

[22] William D. Brooks, "Judging Bias in Intercollegiate Debate," Journal of the American Forensics Association, VII (Winter, 1971), 197-200.

[23] Brooks, p. 198.

[24] The one tournament in which this did not occur was the Illinois State University "Tournament of Champions," a stringently limited tournament in which teams are drawn from throughout the nation through a difficult qualification procedure. Moreover, as a result of this qualification procedure, the calibre of competition and judging at this tournament tends to be well above the average of other debate tournaments.

[25] Michael T. Hayes and Joe McAdoo, "Debate Performance: Differences Between Male and Female Rankings," Journal of the American Forensics Association, VIII (Winter, 1972), 127-131.

[26] Wayne E. Hensley and David B. Strother, "Success in Debate," The Speech Teacher, XVII (September, 1968), 235-237.

[27] This fact is so widely known and accepted as to require little documentation. If documentation is desired, the reader is referred to any basic text in argumentation and debate. A good example is James H. McBurney and Glen E. Mills, Argumentation and Debate: Techniques of a Free Society (New York: The Macmillan Company, 1964), p. 318.

[28] Hayes and McAdoo, p. 127.

[29] Hensley and Strother, p. 236.

[30] This is an assertion for which there is no available evidence. It is based on the author's nine years of experience in intercollegiate debate. There are any number of studies which suggest, indirectly, that some concern over prestige does exist among students and coaches, but the variable has never been examined experimentally. For inferences concerning the possible existance of such a variable, see the study mentioned above by Ellis and Minter.

[31] Brooks, pp. 199-200.

# CHAPTER II

# METHODOLOGY

## Sampling Procedures

In order to eliminate possible effects due to a specific debate topic, the data were collected from tournaments covering a five year period.[1] The experimental samples were stratified between the five years, and each stratum was stratified between debates occuring in the first half of the tournament, those occuring in the second half, and those coming during the elimination rounds. In both cases, the weights were equalized for each stratum. This procedure was adopted to eliminate possible effects across time, either between or within tournaments.

It was decided to accept a boundary on the error of estimation of no more than one point on the 5-30 point scale for speaker ratings in the estimation of sample means. In order to achieve that limit, the following formula for determining the size of the samples was used.[2]

$$n = \frac{\sum_{i=1}^{L} \frac{N_i^2 \sigma_i^2}{w_i}}{N^2 \frac{B^2}{4} + \sum_{i=1}^{L} N_i \sigma_i^2}$$

where: $N_i$ is the total number of debates in stratum i;

$\sigma_i^2$ is the population variance of stratum i;

N is the total number of debates; and

$w_i$ is the fraction of observations allocated to stratum i.

Since this formula requires some estimate of population variance for each of the sample strata, a 1 in 20 systematic sample with N = 200 was drawn from the data pool to provide those estimates.

After calculation of the variance for each of the strata, the appropriate sample size was computed to be 75. Samples for the study were drawn from the data pool by a 1 in 30 systematic sample until each stratum was filled. The allocation of the total sample across each stratum was derived by the formula $n_i = nw_i$.

Dependent Measures

Examination of the data began with a study of the four dependent measures available from the A.F.A. Form "C" ballot; win-loss decision, speaker rating, team rating, and speaker ranking. Since all four measures are reports of the same event (i.e. one debate), the assumption was that there would be a relationship among them. The exact form of that relationship, and its strength, were examined to determine if a more parsimonious statistical procedure could be employed.

Specifically, if a strong positive correlation existed among the variables speaker ranking, speaker rating, team rating, and winning, the testing of the research hypotheses might be collapsed to examine that one of the dependent measures which accounted for the greatest portion of the variance. Inferences concerning those dependent measures not tested directly could still be made on the basis of the relationships known to exist.

In order to determine the relationship among the available measures of outcome for an intercollegiate debate: win-loss decision; speaker rankings; speaker ratings; and team ratings; the data were submitted to both correlation and regression analysis.

To facilitate sampling, the data pool was arranged in consecutive order of rounds and chronological order of the tournaments. Two samples, each with $N = 200$, were selected from the data pool through a 1 in 20 systematic sample. In instances where this procedure produced an incomplete ballot, the selection moved backwards one step at a time until a complete ballot was encountered. In each case, the progression continued from the ballot actually included in the sample. In no instance did this correction procedure require moving more than three ballots away from that indicated by the 1 in 20 progression.

The data from the first sample were submitted to analysis by the BMD02D program,[3] and produced the correlation matrix shown in Table I.

TABLE I

Pearson Product-Moment Correlations for Dependent Measures of Outcome

| | Winning | Losing | Speaker Rating | Team Rating | Speaker Ranking |
|---|---|---|---|---|---|
| Winning | 1.0000 | -1.0000 | .3842 | .2902 | -.6243 |
| Losing | | 1.0000 | -.3842 | -.2902 | .6243 |
| Speaker rating | | | 1.0000 | .7592 | -.4646 |
| Team rating | | | | 1.0000 | -.2120 |
| Speaker ranking | | | | | 1.0000 |

The correlation coefficients were tested for significant deviation from zero using a form of the "t" test adapted for the standard deviation of "r", the estimated value of the population correlation coefficient.[4] The calculated value of the test statistic "t" for each correlation coefficient and the interpretation associated with that statistic are shown in Table II.

TABLE II

Tests for Significance of Pearson Product-Moment Correlations

| Variables | $r_{ij}$ | t-value | interpretation |
|---|---|---|---|
| Win/loss | -1.0000 | undefined | significant with $p<\infty$ |
| Win/Sp. rating | .3842 | 18.43 | significant with $p<.01$ |
| Win/Team rating | .2902 | 13.93 | significant with $p<.01$ |
| Win/Sp. ranking | -.6243 | -35.46 | significant with $p<.01$ |
| Lose/Sp. rating | -.3842 | -18.43 | significant with $p<.01$ |
| Lose/team rating | -.2902 | -13.93 | significant with $p<.01$ |
| Lose/Sp. ranking | .6243 | 35.46 | significant with $p<.01$ |
| Sp. rating/team rating | .7592 | 51.78 | significant with $p<.01$ |
| Sp. rating/Sp. ranking | -.4646 | -23.36 | significant with $p<.01$ |
| Team rating/Sp. rank | -.2120 | -12.66 | significant with $p<.01$ |

The results of this analysis conformed to the a priori expectation of significant correlations based on the fact that all four dependent variables are reports of the same event. The absolute negative correlation between variables 1 and 2 was, of course, forced by the mutually exclusive definitions of those two nominal variables. The negative correlations between speaker ratings and team ratings and speaker rankings is explained by the opposing directions of the scales on which those variables are measured. As performance improves, speaker and team ratings move upward on scales of 5-30 and 1-5 points respectively, whereas speaker rank moves downward toward 1 on a scale of 1 to 4.

The finding of a significant relationship between the four available dependent measures suggested the presence of a principal component effect among them. If true, that principal component would serve as a common "index of outcome" which would pool the information contained in the four dependent measures. The search for a principal component was begun using the BMDX72 "Factor Analysis" program.[5] The data from the second sample were used for this analysis, with N = 200.

Factor analysis revealed that the measures of outcome available from the A.F.A. Form "C" ballot clustered around four factors which, taken together, accounted for 100% of the variance in the original four dependent measures.

However, the proportion of variance accounted for was not equally distributed among the four factors. A principal component was discovered which accounted for some 60 percent of the total variance. A second factor, which accounted for an additional 25 percent of the variance, also emerged.

The output of BMDX72 also included a definition of both factors in terms of their linear relationship to the four original measures. Those definitions were as follows.

Factor one: $y1 = .73x1 - .37x2 - .19x3 + 1.30x4$

Factor two: $y2 = .52x1 + .39x2 + .61x3 + .49x4$

where x1 was a dichotomous variable representing win-loss;

x2 was speaker rating;

x3 was team rating; and

x4 was speaker ranking.

On the basis of this analysis, it was concluded that the use of Factor one as the "index of outcome" would be of interest. Because it represents contributions from each of the four dependent measures, the factor would itself contain more information than any one of the four.

In all subsequent analyses, the data were first submitted to the BMDX72 program in order to generate "index of outcome" scores for each case. Those scores were then used as an additional response measure for testing regression by the dependent variables.

Predictor Variables

The method chosen for the testing of the research hypotheses was an adaptation of multi-variate regression analysis, using "dummy" variables to introduce nominal data into the regression equation. Since the variables to be tested all lay outside the range of "ability" variables as defined in Chapter I, any significant apportioning of variance in the dependent measure to a variable under study was taken as evidence of a non-ability effect.

Regression analysis was chosen as the appropriate statistical procedure because the issue of importance was the estimation of relationships among variables influencing the outcome of intercollegiate debates rather than measuring the strength of those relationships.

Actual calculations were performed by the University of Florida Computing Center using the BMDX63 program "Multivariate General Linear Hypothesis" developed by the U.C.L.A. Health Sciences Computing Facility. The output from this program includes regression coefficients for each of the predictor variables in the model, various cross-product matrices, and appropriate "F" statistics with associated degrees of freedom for hypotheses selected by the experimenter.[6]

The procedure required a division of the study into two steps. Attention was first given to the format variables,"side of topic" and "speaker position," in

order to determine whether or not blocking for the effects of those variables might be appropriate when examining the rater variables "sex," "prestige," and "proximity."

The general form of the model tested was

$$y = \beta_o + \beta_1 x_1 + \ldots + \beta_k x_k + \varepsilon, \text{ where}$$

y was the index of outcome;

$\beta_o$ was some constant;

$x_1 \ldots x_k$ were the non-ability variables being studied;

$\beta_1 \ldots \beta_k$ were the regression weights associated with each variable; and

$\varepsilon$ represented unapportioned variance, or error in the model.

## Statistical Models

Examination of the research hypotheses dealing with the effects of format bias on the dependent measure index of outcome utilized the following model.

Model #1.

$$\hat{y} = b_o + b_1 x_1 + b_2 x_2 + b_3 x_3 + \varepsilon, \text{ where}$$

$\hat{y}$ is the dependent variable, index of outcome;

$b_o$ is the mean score for 2nd negative speakers;

$b_1$ is the mean score for 1st affirmative speakers, minus $b_o$;

$b_2$ is the mean score for 2nd affirmative speakers, minus $b_o$;

$b_3$ is the mean score for 1st negative speakers, minus $b_o$;

$x_1$ is a dummy variable for 1st affirmatives (1 if 1st affirmative, 0 otherwise);

$x_2$ is a dummy variable for 2nd affirmatives (1 if 2nd affirmative, 0 otherwise); and

$x_3$ is a dummy variable for 1st negatives (1 if 1st negative, 0 otherwise).

The model was tested for "goodness of fit" using the test statistic

$$F = \frac{MS(\text{Lack of Fit})}{MS(\text{Experimental Error})} \quad F_{p-(k+1),\, n-p}$$

using the rejection region

$$F > F_{.05,\, p-(k+1),\, n-p}$$

Having verified that the derived model did in fact provide a good fit for the sample data, the hypotheses were tested specifically as follows.

Hypothesis 1.2: There is no effect due to side in the individual speaker scores in intercollegiate debate.

$$H_o: \beta_1 = \beta_2 = \beta_3 = 0$$

$$H_a: \text{at least one } \beta_i \neq 0$$

Hypothesis 2.1: There is no effect due to speaker position in the scores of affirmative debaters.

$$H_o: \beta_1 = 0 \qquad H_o: \beta_2 = 0$$

$$H_a: \beta_1 \neq 0 \qquad H_a: \beta_2 \neq 0$$

Hypothesis 2.2: There is no effect due to speaker position in the scores of negative debaters.

$$H_o: \beta_3 = 0$$
$$H_a: \beta_3 \neq 0$$

All tests were made using a preset value of $\alpha = .05$ and the test statistic

$$F = \frac{MSR}{MSE} \quad F_{\nu_1 \nu_2}$$ [7]

As the examination of the appropriate research hypotheses provided no evidence of an effect on the outcome of debates due to format bias, analysis of the data was continued by examining the effects due to the three rater bias variables "sex," "prestige," and "proximity." The examination of research hypothesis three, dealing with the effect of the sex of the debaters and the judge on the index of outcome, was based on the following model.

Model #2.

$$\hat{y} = b_o + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + \varepsilon$$

where $\hat{y}$ is the dependent variable, index of outcome;

$b_o$ is the mean score for a male debater with a male colleague evaluated by a male judge;

$b_1$ is the mean score for a male debater with a female colleague evaluated by a male judge, minus $b_o$;

$b_2$ is the mean score for a male debater with a male colleague evaluated by a female judge, minus $b_o$;

$b_3$ is the mean score for a male debater with a female colleague evaluated by a female judge, minus $b_o$;

$b_4$ is the mean score for a female debater with a female colleague, evaluated by a male judge, minus $b_o$;

$b_5$ is the mean score for a female debater with a female colleague, evaluated by a female judge, minus $b_o$;

$x_1$ is a dummy variable (1 if male debater with female colleague and male judge, 0 otherwise);

$x_2$ is a dummy variable (1 if male debater with male colleague and female judge, 0 otherwise);

$x_3$ is a dummy variable (1 if male debater with female colleague and female judge, 0 otherwise);

$x_4$ is a dummy variable (1 if female debater with female colleague and male judge, 0 otherwise); and

$x_5$ is a dummy variable (1 if female debater with female colleague and female judge, 0 otherwise).

Having verified that the derived model did in fact provide a good fit for the sample data, the hypotheses were tested specifically as follows. The general hypothesis of some effect due to sex was of the form

$$H_o: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

$$H_a: \text{at least one } \beta_i \neq 0.$$

In order to isolate the specific item in which an effect due to sex was to be found, each of the terms in

the model was tested under the general form

$$H_o: \beta_i = 0$$

$$H_a: \beta_i \neq 0$$

All tests were made using a preset value of $\alpha = .05$ and the test statistic

$$F = \frac{MSR}{MSE} \quad F_{\nu_1 \nu_2} \quad (8)$$

The examination of research hypothesis four, dealing with the effects of prestige on the index of outcome, was based on the following model.

Model #3.

$$\hat{y} = b_o + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + b_8x_8 + \varepsilon, \text{ where}$$

$\hat{y}$ is the dependent variable;

$b_o$ is the mean score for a high prestige team debating a low prestige team with a judge from a high prestige school;

$b_1$ is the mean score for a high prestige team debating a low prestige team with a judge from a low prestige school, minus $b_o$;

$b_2$ is the mean score for a high prestige team debating a low prestige team with a judge from a junior college, minus $b_o$;

$b_3$ is the mean score for a high prestige team debating a junior college with a judge from a high prestige school, minus $b_o$;

$b_4$ is the mean score for a high prestige team debating a junior college with a judge from a low prestige school, minus $b_o$;

$b_5$ is the mean score for a high prestige team debating a junior college with a judge from a junior college, minus $b_o$;

$b_6$ is the mean score for a low prestige team debating a junior college with a judge from a high prestige school, minus $b_o$;

$b_7$ is the mean score for a low prestige team debating a junior college with a judge from a low prestige school, minus $b_o$;

$b_8$ is the mean score for a low prestige team debating a junior college with a judge from a junior college, minus $b_o$;

$x_1$ is a dummy variable (1 if high/low with low judge, 0 otherwise);

$x_2$ is a dummy variable (1 if high/low with J.C. judge, 0 otherwise);

$x_3$ is a dummy variable (1 if high/j.c. with high judge, 0 otherwise);

$x_4$ is a dummy variable (1 if high/j.c. with low judge, 0 otherwise);

$x_5$ is a dummy variable (1 if high/j.c. with j.c. judge, 0 otherwise);

$x_6$ is a dummy variable (1 if low/j.c. with high judge, 0 otherwise);

$x_7$ is a dummy variable (1 if low/j.c. with low judge, 0 otherwise); and

$x_8$ is a dummy variable (1 if low/j.c. with j.c. judge, 0 otherwise).

Having verified that the derived model did in fact provide a good fit for the sample data, the hypotheses were tested specifically as follows. The general hypothesis of some effect due to prestige was of the form

$H_o$: $\beta_1 = \beta_2 = \beta_3 + \beta_4 = \beta_5 + \beta_6 + \beta_7 = \beta_8 = 0$

$H_a$: at least one $\beta_i \neq 0$.

In order to isolate the specific item in which an effect due to prestige was to be found, each of the terms in the model was tested under the general form

$$H_o: \beta_i = 0$$

$$H_a: \beta_i \neq 0.$$

All tests were made using a preset value of $\alpha$ .05 and the test statistic

$$F = \frac{MSR}{MSE} \qquad F_{\nu_1 \nu_2}^{9}$$

The examination of research hypothesis five, dealing with the effects of "proximity" on the index of outcome, was based on the following model.

Model #4.

$$\hat{y} = b_o + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + \varepsilon, \text{ where}$$

$\hat{y}$ is the dependent variable;

$b_o$ is the mean score for a debater, with opponents and judge from his own district, who is nearest to the judge;

$b_1$ is the mean score for a debater, with opponents and judge from his own district, who is farthest from the judge, minus $b_o$;

$b_2$ is the mean score for a debater, with opponents from his own district and a judge from another district, who is nearest to the judge, minus $b_o$;

$b_3$ is the mean score for a debater, with opponents from his own district and a judge from another district, who is farthest from the judge, minus $b_o$;

$b_4$ is the mean score for a debater, with opponents and the judge each from different districts, who is nearest to the judge, minus $b_o$;

$b_5$ is the mean score for a debater, with opponents and the judge each from different districts, who is farthest from the judge, minus $b_o$;

$b_6$ is the mean score for a debater with a judge from his own district and opponents from another district, minus $b_o$;

$b_7$ is the mean score for a debater with opponents from another district and the judge from his opponent's district, minus $b_o$;

$x_1$ is a dummy variable (1 if of the conditions for $b_1$, 0 otherwise);

$x_2$ is a dummy variable (1 if of the conditions for $b_2$, 0 otherwise);

$x_3$ is a dummy variable (1 if of the conditions for $b_3$, 0 otherwise);

$x_4$ is a dummy variable (1 if of the conditions for $b_4$, 0 otherwise);

$x_5$ is a dummy variable (1 if of the conditions for $b_5$, 0 otherwise);

$x_6$ is a dummy variable (1 if of the conditions for $b_6$, 0 otherwise);

$x_7$ is a dummy variable (1 if of the conditions for $b_7$, 0 otherwise).

Having verified that the derived model did in fact provide a good fit for the sample data, the hypotheses were tested specifically as follows. The general hypothesis of some effect due to proximity was of the form

$H_o$: $\beta_1 + \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$

$H_a$: at least one $\beta_i \neq 0$.

In order to isolate the specific item in which an effect due to prestige was to be found, each of the terms in the model was tested under the general form

$H_o$: $\beta_i = 0$

$H_a$: $\beta_i \neq 0$.

All tests were made using a preset value of $\alpha = .05$ and the test statistic

$$F = \frac{MSR}{MSE} \qquad F_{\nu_1 \nu_2}$$

10

## NOTES - CHAPTER II

[1] The years covered were the academic years 1967-68, 1968-69, 1969-70, 1970-71, and 1971-72. The tournaments involved were the Peachtree Debate Tournament, hosted by Emory University in Atlanta, Georgia; the Birmingham Invitational Debate Tournament, hosted by Samford University in Birmingham, Alabama; and the Gator Invitational Debate Tournament, hosted by the University of Florida in Gainesville, Florida.

[2] William Mendenhall, Lyman Ott, and Richard Schaeffer, Elementary Survey Sampling (Belmont, California: Wadsworth Publishing Company, Inc., 1971), p. 61.

[3] For a description of this program, see W.J. Dixon, BMD Biomedical Computer Programs (University of California Publications in Automatic Computation No. 2, 1972), pp. 49-59.

[4] The test statistic took the form:

$$t = \frac{r_{ij}}{\sqrt{1-r_{ij}^2}} \sqrt{n-2}$$

For a discussion of this test, see William Hayes, Statistics (New York: Holt, Rinehart and Winston, 1968), pp. 529-533.

[5] For a description of this program, see W.J. Dixon, BMD Biomedical Computer Programs: X-series Supplement (University of California Publications in Automatic Computation No. 3, 1972), pp. 90-103.

[6] For a more complete description of this program, see Dixon, X-series Supplement, pp. 23-33.

[7] Actual calculation of these values was performed through the BMDX63 program.

[8] Actual calculation of these values was performed through the BMDX63 program.

[9] Actual calculation of these values was performed through the BMDX63 program.

[10] Actual calculation of these values was performed through the BMDX63 program.

# CHAPTER III

## RESULTS

The format variables "side of topic" and "speaker position" were the first to be studied, using Model #1 as described in Chapter II. Examination of comparative wins for affirmative and negative debate teams indicated no significant difference in outcome due to the side of the topic. The appropriate regression weights for Model #1, as calculated by BMDX63, and the results of individual tests for significant regression, are shown in Table III.

TABLE III

Regression Weights for Format Variables

| Variable | Weight | F-score | d.f. | Interpretation |
|---|---|---|---|---|
| x1 | .068 | 0.009 | 1,76 | $p > .05$ |
| x2 | .303 | 0.179 | 1,76 | $p > .05$ |
| x3 | .913 | 1.629 | 1,76 | $p > .05$ |

The data provide no evidence of an effect on the outcome of intercollegiate debates due to the side of topic or speaker positions. For this reason, analysis of the rater bias variables did not employ blocking along format variables.

Further analysis of these data focused on the regression effects of side of topic and speaker position on each of the four dependent measures provided by the A.F.A. Form "C" ballot. Only in one instance did the results indicate a significant effect. The calculated regression weights for each variable and the F statistic associated with each are shown in Table IV.

The presence of regression on speaker rank indicates that the appropriate formula for predicting an individuals speaker rank would be

$$\hat{y} = 2.0 + .75x, \quad \text{where}$$

$\hat{y}$ is the expected rank and x is a dichotomous variable representing the second affirmative speaker.

Blocking across time within a tournament also produced significant differences. The frequency of wins in preliminary rounds appeared to differ from that in elimination rounds, with a calculated value of F significant with $p < .05$. A Chi-square analysis of the data to determine exact differences revealed that, when considering only preliminary rounds, negative teams won

TABLE IV

Regression Weights for Format Variables On Ballot Measures

Win-loss

| Variable | Weight | F-score | d.f. |
|---|---|---|---|
| x1 | -0.30 | 3.76 | 1,76 |
| x2 | -0.30 | 3.76 | 1,76 |
| x3 | 0.00 | 0.00 | 1,76 |

Speaker Rating

| | | | |
|---|---|---|---|
| x1 | -2.25 | 2.54 | 1,76 |
| x2 | -1.70 | 1.45 | 1,76 |
| x3 | .0.80 | 0.32 | 1,76 |

Team Rating

| | | | |
|---|---|---|---|
| x1 | -0.15 | 0.28 | 1,76 |
| x2 | -0.15 | 0.28 | 1,76 |
| x3 | 0.00 | 0.00 | 1,76 |

Speaker Rank

| | | | |
|---|---|---|---|
| x1 | .70 | 3.92 | 1,76 |
| x2 | .75 | 4.51 | 1,76* |
| x3 | .45 | 1.62 | 1,76 |

* significant with $p < .05$

more debates. Chi-square for this comparison was significant with $p < .10$. Negative teams also won more elimination round debates, but the probability of error in this statistic exceeded .10. The percentages by which negative wins exceeded expected frequencies for both preliminary and elimination rounds are shown in Table V.

TABLE V

Percentages of Excessive Negative Wins

| | Preliminary Rounds | Elimination Rounds |
|---|---|---|
| 1967-68 | 50% | 100% |
| 1968-69 | 40% | 25% |
| 1969-70 | 40% | 25% |
| 1970-71 | 25% | 25% |
| 1971-72 | 40% | 25% |
| Totals | 27% | 68% |

Sex

The effect of the rater variable "sex" on the index of outcome was examined with Model #2, as described in Chapter II. The predictors for this analysis were

dichotomous variables representing the nine possible combinations of sex when considering a debater, his colleague, and the judge. The calculated regression weights, F-scores, and interpretations for this model are shown in Table VI.

TABLE VI

Regression Weights for Rater Variable Sex

| Variable | Weight | F-score | d.f. | Interpretation |
|---|---|---|---|---|
| x1 | -0.75 | 0.72 | 1,92 | p > .05 |
| x2 | -1.03 | 0.68 | 1,92 | p > .05 |
| x3 | -0.26 | 0.05 | 1,92 | p > .05 |
| x4 | -1.23 | 0.97 | 1,92 | p > .05 |
| x5 | -0.02 | 0.00 | 1,92 | p > .05 |
| x6 | -1.25 | 1.31 | 1,92 | p > .05 |
| x7 | -0.53 | 0.18 | 1,92 | p > .05 |

These data provide no significant evidence of any effect on outcome due to sex. Analysis then focused on effects of sex on each of the four dependent measures provided on the A.F.A. ballot.

When considering win-loss as the dependent measure, there were a number of significant differences revealed by the data. Sex affected win-loss both for debaters and for judges. These results are shown in Table VII.

TABLE VII

Regression By Sex on Win-Loss for Groups

| Group | F-score | d.f. | Interpretation |
|---|---|---|---|
| Male debaters | 10.37 | 3,92 | $p < .01$ |
| Female debaters | 10.25 | 4,92 | $p < .01$ |
| Male judges | 16.13 | 3,92 | $p < .01$ |
| Female judges | 5.22 | 4,92 | $p < .01$ |

Interaction between the sex of the debater, the sex of his colleague, and the sex of the judge was also considered. In all cases, the calculated regression weights were found to be significant with $p < .05$. These results are shown in Table VIII.

When considering the effects of sex on the dependent measure speaker rating, analysis again indicated the presence of significant regression. No significant effects were found for groups in speaker rating. However, this was not the case when considering interaction among the sex of the debater, the sex of his colleague, and the sex of the judge, These results are shown in Table IX.

TABLE VIII

Regression of Sex on Win-Loss

| Variable | (debater/colleague/judge) | Weight | F-score |
|---|---|---|---|
| x1 | male/male/female | -0.54 | 18.47 |
| x2 | male/female/male | -0.88 | 24.66 |
| x3 | male/female/female | -0.44 | 6.91 |
| x4 | female/male/male | -0.88 | 24.66 |
| x5 | female/male/female | -0.44 | 6.91 |
| x6 | female/female/male | -0.83 | 23.36 |
| x7 | female/female/female | -0.50 | 8.05 |

TABLE IX

Regression of Sex on Speaker Ratings

| Variable | (debater/colleague/judge) | Weight | F-score | Interpretation |
|---|---|---|---|---|
| x1 | male/male/female | -0.30 | 0.08 | $p > .05$ |
| x2 | male/female/male | -3.30 | 4.94 | $p < .05$ |
| x3 | male/female/female | -1.47 | 1.06 | $p > .05$ |
| x4 | female/male/male | -3.68 | 6.12 | $p < .05$ |
| x5 | female/male/female | -3.47 | 5.92 | $p < .05$ |
| x6 | female/female/male | 0.28 | 0.05 | $p > .05$ |
| x7 | female/female/female | -2.30 | 2.40 | $p > .05$ |

Sex affected team ratings both for debaters and by judges. Regression of sex on team ratings by groups is summarized in Table X.

TABLE X

Regression of Sex on Team Ratings for Groups

| Group | F-score | d.f. | Interpretation |
|---|---|---|---|
| Male debaters | 2.34 | 3,92 | $p > .05$ |
| Female debaters | 2.67 | 4,92 | $p < .05$ |
| Male judges | 3.44 | 3,92 | $p < .05$ |
| Female judges | 1.82 | 4,92 | $p > .05$ |

Once again, significant interaction was found to exist among the sex of the debater, the sex of the colleague, and the sex of the judge. The regression weights for various sex combinations and the results of the tests for significance are shown in Table XI.

When considering the effects of sex on the dependent measure speaker ranking, analysis again indicated the presence of significant regression. The differences between groups were tested, and the results are shown in Table XII.

As was the case for the other three dependent measures, speaker ranking revealed interaction effects

TABLE XI

Regression of Sex on Team Ratings

| Variable | (debater/colleague/judge) | Weight | F-score | Interpretation |
|---|---|---|---|---|
| x1 | male/male/female | -0.09 | 0.19 | $p > .05$ |
| x2 | male/female/male | -0.58 | 4.53 | $p < .05$ |
| x3 | male/female/female | -0.48 | 3.40 | $p > .05$ |
| x4 | female/male/male | -0.58 | 4.53 | $p < .05$ |
| x5 | female/male/female | -0.48 | 3.40 | $p > .05$ |
| x6 | female/female/female | 0.13 | 0.32 | $p > .05$ |
| x7 | female/female/female | -0.45 | 2.77 | $p > .05$ |

TABLE XII

Regression of Sex on Speaker Rankings for Groups

| Groups | F-score | d.f. | Interpretation |
|---|---|---|---|
| Male debaters | 4.09 | 3,92 | $p < .05$ |
| Female debaters | 4.26 | 4,92 | $p < .01$ |
| Male judges | 6.60 | 3,92 | $p < .01$ |
| Female judges | 2.19 | 4,92 | $p > .05$ |

between the sex of the debaters and the sex of the judge. The calculated regression weights, with their associated F statistics and interpretations, are shown in Table XIII.

TABLE XIII

Regression of Sex on Speaker Rankings

| Variable | (debater/colleague/judge) | Weight | F-score | Interpretation |
|---|---|---|---|---|
| x1 | male/male/female | 0.80 | 6.49 | $p < .05$ |
| x2 | male/female/male | 1.43 | 10.42 | $p < .01$ |
| x3 | male/female/female | 0.63 | 2.24 | $p > .05$ |
| x4 | female/male/male | 1.43 | 10.42 | $p < .01$ |
| x5 | female/male/female | 0.96 | 5.21 | $p < .05$ |
| x6 | female/female/male | 1.30 | 11.38 | $p < .01$ |
| x7 | female/female/female | 0.80 | 3.28 | $p > .05$ |

Prestige

The effect of the rater variable "prestige" on the index of outcome was examined with Model #3, as described in Chapter II. The predictors for this analysis were dichotomous variables representing the eighteen possible combinations of high prestige, low prestige, and junior college debate teams and judges. The appropriate regression weights for Model #3, as calculated by

BMDX63, and the results of individual tests for significant regression, are shown in Table XIV.

Since none of the above items reached a probability level of .95 or greater, the model was re-evaluated. On the basis of this evidence, the expected result for teams at all prestige levels would be that team's average index of outcome in past debates.

Confirmation of these results was sought by examining the regression of prestige on each of the four dependent measures provided by the A.F.A. Form "C" ballot.

For the dependent measure win-loss, significant regression effects were found for both low prestige and junior college debate teams. In both instances, the calculated F statistic was found to be significant with $p < .01$. When interaction between the prestige of the teams in a given round and the prestige of the judge were considered, a number of significant differences were found to exist. These results are shown in Table XV.

When considering the dependent measure speaker rating, the prestige of a debate team was found to have a specific effect with $p < .05$. Significant regression effects were found for debate teams at all prestige levels, and for the scores given by judges of all prestige levels. In each instance, the calculated F statistic was found to be significant with $p < .05$.

TABLE XIV

Regression Weights for Rater Variable Prestige

| Variable | Weight | F-score | d.f. | Interpretation |
|---|---|---|---|---|
| x1 | 0.21 | .015 | 1,102 | p > .05 |
| x2 | -1.05 | .331 | 1,102 | p > .05 |
| x3 | 1.54 | .802 | 1,102 | p > .05 |
| x4 | 1.10 | .410 | 1,102 | p > .05 |
| x5 | 0.32 | .034 | 1,102 | p > .05 |
| x6 | 0.39 | .034 | 1,102 | p > .05 |
| x7 | 1.07 | .413 | 1,102 | p > .05 |
| x8 | 1.52 | .702 | 1,102 | p > .05 |
| x9 | 0.83 | .268 | 1,102 | p > .05 |
| x10 | 0.93 | .316 | 1,102 | p > .05 |
| x11 | -0.12 | .006 | 1,102 | p > .05 |
| x12 | 0.11 | .004 | 1,102 | p > .05 |
| x13 | 1.90 | 1.221 | 1,102 | p > .05 |
| x14 | 2.01 | 1.371 | 1,102 | p > .05 |
| x15 | 1.58 | .972 | 1,102 | p > .05 |
| x16 | 1.07 | .418 | 1,102 | p > .05 |
| x17 | 1.43 | .472 | 1,102 | p > .05 |

TABLE XV

Regression of Prestige on Win-Loss

| Variable | (team/opponent/judge) | Weight | F-score | Interpretation |
|---|---|---|---|---|
| x1 | high/low/low | -0.25 | 0.58 | p > .05 |
| x2 | high/low/jr.coll. | -0.50 | 1.93 | p > .05 |
| x3 | high/jr.coll./high | 0.00 | 0.00 | p > .05 |
| x4 | high/jr.coll./low | -0.33 | 0.96 | p > .05 |
| x5 | high/jr.coll./jr.coll. | -0.67 | 3.85 | p > .05 |
| x6 | low/hihg/high | -1.00 | 5.77 | p < .05 |
| x7 | low/high/low | -0.75 | 5.20 | p < .05 |
| x8 | low/high/jr.coll. | -0.50 | 1.93 | p > .05 |
| x9 | low/jr.coll./high | -0.17 | 0.28 | p > .05 |
| x10 | low/jr.coll./low | -0.50 | 2.31 | p > .05 |
| x11 | low/jr.coll./jr.coll. | -1.00 | 9.24 | p < .01 |
| x12 | jr.coll./high/high | -1.00 | 8.66 | p < .01 |
| x13 | jr.coll./high/low | -0.67 | 3.85 | p > .05 |
| x14 | jr.coll./high/jr.coll. | -0.33 | 0.96 | p > .05 |
| x15 | jr.coll./low/high | -0.83 | 6.87 | p < .05 |
| x16 | jr.coll./low/low | -0.50 | 2.31 | p > .05 |
| x17 | jr.coll./low/jr.coll. | -0.00 | 0.00 | p > .05 |

TABLE XVI

Regression of Prestige on Speaker Rating

| Variable | (team/opponent/judge) | Weight | F-score | Interpretation |
|---|---|---|---|---|
| x1 | high/low/low | -9.00 | 8.84 | $p < .01$ |
| x2 | high/low/jr.coll. | -3.50 | 1.11 | $p > .05$ |
| x3 | high/jr.coll./high | -3.17 | 1.03 | $p > .05$ |
| x4 | high/jr.coll./low | -5.67 | 3.28 | $p > .05$ |
| x5 | high/jr.coll./jr.coll. | -7.17 | 5.25 | $p < .05$ |
| x6 | low/high/high | -4.50 | 1.38 | $p > .05$ |
| x7 | low/high/low | -10.50 | 12.03 | $p < .01$ |
| x8 | low/high/jr.coll. | -4.00 | 1.45 | $p > .05$ |
| x9 | low/jr.coll./high | -5.08 | 3.02 | $p > .05$ |
| x10 | low/jr.coll./low | -8.63 | 8.11 | $p < .01$ |
| x11 | low/jr.coll./jr.coll. | -10.50 | 12.03 | $p < .01$ |
| x12 | jr.coll./high/high | -5.50 | 3.09 | $p > .05$ |
| x13 | jr.coll./high/low | -6.83 | 4.78 | $p < .05$ |
| x14 | jr.coll./high/jr.coll. | -7.83 | 6.28 | $p < .05$ |
| x15 | jr.coll./low/high | -10.00 | 11.69 | $p < .01$ |
| x16 | jr.coll./low/low | -8.13 | 7.20 | $p < .01$ |
| x17 | jr.coll./low/jr.coll. | -9.13 | 9.08 | $p < .01$ |

When interaction between the prestige of the teams in a given round and the prestige of the judge was considered, a number of significant differences were found to exist. These results are shown in Table XVI.

When considering the dependent measure team rating, analysis revealed significant differences both among expected scores by teams and from judges at the different levels of prestige. These results are summarized in Table XVII.

TABLE XVII

Regression of Prestige on Team Ratings for Groups

| Group | F-score | d.f. | Interpretation |
|---|---|---|---|
| High prestige teams | 2.88 | 5,102 | $p < .05$ |
| Low prestige teams | 3.81 | 6,102 | $p < .01$ |
| Junior college teams | 4.38 | 6,102 | $p < .01$ |
| High prestige judges | 5.10 | 5,102 | $p < .01$ |
| Low prestige judges | 2.76 | 6,102 | $p < .05$ |
| Junior college judges | 4.25 | 6,102 | $p < .01$ |

Interaction between the prestige of the two teams in the debate round and the prestige of the judge was also tested. Again, significant differences in the expected team ratings were discovered. The calculated regression

weight for each combination, along with its F statistic and the interpretation of that score, is shown in Table XVIII.

When considering the dependent measure speaker ranking, analysis revealed significant difference only for the group of low prestige debate teams. The F statistic calculated for this group by BMDX63 was found to be significant with $p < .05$. No significant differences were found for teams at other prestige levels, or for the scores given by judges at any prestige level. However, when interaction between the prestige levels of the debate teams and the prestige level of the judge was considered, a number of significant differences were revealed. The regression weights for each combination, along with the appropriate F statistic and its interpretation, are shown in Table XIX.

Proximity

The effect of the rater variable "proximity" on the index of outcome was examined with Model #4, as described in Chapter II. The predictors for this analysis were dichotomous variables representing the eight possible combinations of teams and judges when considering both geographical proximity and N.D.T. district memberships. The appropriate regression weights for Model #4, as calculated by BMDX63, and the results of individual tests for significant regression, are shown in Table XX.

TABLE XVIII

Regression of Prestige on Team Ratings

| Variable | (team/opponents/judge) | Weight | F-score | Interpretation |
|---|---|---|---|---|
| x1 | high/low/low | -1.75 | 7.99 | $p < .01$ |
| x2 | high/low/jr.coll. | -0.50 | 0.54 | $p > .05$ |
| x3 | high/jr.coll./high | -1.00 | 2.45 | $p < .05$ |
| x4 | high/jr.coll./low | -1.33 | 4.35 | $p < .01$ |
| x5 | high/jr.coll./jr.coll. | -1.67 | 6.80 | $p < .01$ |
| x6 | low/high/high | -1.00 | 1.63 | $p > .05$ |
| x7 | low/high/low | -2.00 | 10.45 | $p < .01$ |
| x8 | low/high/jr.coll. | -1.00 | 2.18 | $p > .05$ |
| x9 | low/jr.coll./high | -1.33 | 4.97 | $p < .01$ |
| x10 | low/jr.coll./low | -2.00 | 10.45 | $p < .01$ |
| x11 | low/jr.coll./jr.coll. | -2.25 | 13.33 | $p < .01$ |
| x12 | jr.coll./high/high | -1.33 | 4.35 | $p < .01$ |
| x13 | jr.coll./high/low | -2.33 | 13.33 | $p < .01$ |
| x14 | jr.coll./high/jr.coll. | -2.00 | 9.79 | $p < .01$ |
| x15 | jr.coll./low/high | -2.50 | 17.49 | $p < .01$ |
| x16 | jr.coll./low/low | -1.75 | 7.99 | $p < .01$ |
| x17 | jr.coll./low/jr.coll. | -1.50 | 5.88 | $p < .01$ |

TABLE XIX

Regression of Prestige on Speaker Rankings

| Variable | (team/opponent/judge) | Weight | F-score | Interpretation |
|---|---|---|---|---|
| x1 | high/low/low | 0.25 | 0.09 | $p > .05$ |
| x2 | high/low/jr.coll. | 0.50 | 0.29 | $p > .05$ |
| x3 | high/jr.coll./high | 0.33 | 0.15 | $p > .05$ |
| c4 | high/jr.coll./low | 0.83 | 0.93 | $p > .05$ |
| x5 | high/jr.coll./jr.coll. | 1.00 | 1.34 | $p > .05$ |
| x6 | low/high/high | 2.00 | 3.57 | $p < .01$ |
| x7 | low/high/low | 1.75 | 4.37 | $p < .01$ |
| x8 | low/high/jr.coll. | 1.50 | 2.67 | $p < .05$ |
| x9 | low/jr.coll./high | 0.25 | 0.09 | $p > .05$ |
| x10 | low/jr.coll./low | 0.87 | 1.09 | $p > .05$ |
| x11 | low/jr.coll./jr.coll. | 1.38 | 2.70 | $p < .05$ |
| x12 | jr.coll./high/high | 1.67 | 3.71 | $p < .01$ |
| x13 | jr.coll./high/low | 1.17 | 1.82 | $p > .05$ |
| x14 | jr.coll./high/jr.coll. | 1.00 | 1.34 | $p > .05$ |
| x15 | jr.coll./low/high | 1.75 | 4.68 | $p < .01$ |
| x16 | jr.coll./low/low | 1.13 | 1.81 | $p > .05$ |
| x17 | jr.coll./low/jr.coll. | 0.50 | 0.36 | $p > .05$ |

TABLE XX

Regression Weights for Rater Variable Proximity

| Variable | Weight | F-score | Interpretation |
|---|---|---|---|
| x1 | 2.87 | 6.12 | $p < .05$ |
| x2 | 1.36 | 1.24 | $p > .05$ |
| x3 | 0.32 | 0.07 | $p > .05$ |
| x4 | 1.83 | 2.66 | $p > .05$ |
| x5 | 1.59 | 2.00 | $p > .05$ |
| x6 | 1.77 | 2.51 | $p > .05$ |
| x7 | 1.20 | 1.15 | $p > .05$ |

Further explanation of the mixed results among the above tests was sought by considering the regression of proximity on each of the four dependent measures provided by the A.F.A. Form "C" ballot.

For the dependent measure win-loss, no significant regression effects were discovered, either for geographical proximity or for district membership. Examination of interaction effects also produced no significant differences. In all cases, $p > .05$.

When considering the dependent measure speaker rating, the proximity of a debate team to the judge was found to have a significant effect with $p < .05$. When

interaction between teams' and judges' district memberships was considered, a number of significant differences were found. These results are shown in Table XXI.

TABLE XXI

Regression Of Proximity On Speaker Ratings

| Variable | Weight | F-score | Interpretation |
|---|---|---|---|
| x1 | -0.92 | 0.42 | p > .05 |
| x2 | -1.85 | 1.56 | p > .05 |
| x3 | -4.65 | 11.13 | p < .01 |
| x4 | -0.11 | 0.01 | p > .05 |
| x5 | -1.54 | 1.27 | p > .05 |
| x6 | -1.11 | 0.66 | p > .05 |
| x7 | -1.54 | 1.27 | p > .05 |

The results for tests of the effect of proximity on the dependent measure team rating were similar to those for speaker rating. When interaction effects were tested, a number of significant differences were discovered. The calculated regression weight for each combination, along with its F statistic and the interpretation of that score, is shown in Table XXII.

TABLE XXII

Regression of Proximity on Team Rating

| Variable | Weight | F-score | Interpretation |
|---|---|---|---|
| x1 | -0.33 | 1.60 | $p > .05$ |
| x2 | -0.23 | 0.71 | $p > .05$ |
| x3 | -0.63 | 5.24 | $p < .05$ |
| x4 | -0.26 | 1.06 | $p > .05$ |
| x5 | -0.41 | 2.54 | $p > .05$ |
| x6 | -0.41 | 2.54 | $p > .05$ |
| x7 | -0.26 | 1.06 | $p > .05$ |

Examination of the dependent measure speaker ranking revealed no significant differences between means for the various groups. This trend continued when interaction was considered. In all cases, the calculated regression weights were found not to be significant.

# CHAPTER IV

# CONCLUSIONS

The results of this study indicate that the format variables "side of topic" and "speaker position" have no significant effect on the overall outcome of intercollegiate debates as measured by the dependent variable index of outcome. This result was confirmed when the effect of format variables on the dependent measure win-loss was considered. The thrust of this evidence was to support the validity of the format currently used in intercollegiate debate.

Specifically, hypotheses 1.1, 1.2, 1.3, 2.1, 2.2, and 2.4 were not rejected. In terms of overall performance and of win-loss, there is no evidence to suggest that the format of intercollegiate debating is biased toward either side or any speaker position.

Hypothesis 2.3 was rejected, with $\alpha = .05$. The corrected model for the prediction of speaker rank on the basis of speaker position was as follows.

$$\hat{y} = 2.0 + .75x + \varepsilon,$$

where $\hat{y}$ is the expected rank and x is a dichotomous variable representing the second affirmative speaker.

The rejection of hypothesis 2.3 leads to the conclusion that second affirmative speakers tend to rank lower than speakers in any other position. This raises doubts as to the validity of using speaker rank as a factor in determining the outcome of intercollegiate debates. The use of the measure would unfairly handicap second affirmative speakers.

The rater variable "sex" appears to have a significant effect on some measures of the outcome of debates. In terms of the overall measure index of outcome, there was no observed effect due to sex. This finding disagrees with the conclusions of Hayes and McAdoo and Hensley and Strother.

However, when focusing directly on the dependent measure win-loss, the conclusions of Hensley and Strother were confirmed. Specifically, hypotheses 3.1, 3.4, and 3.7 were rejected with $\alpha = .05$. The corrected form of Model #2 for prediction of win-loss results by sex was al follows.

$$\hat{y} = 1.00 - .54x_1 - .88x_2 - .44x_3 - .88x_4 -$$

$$.44x_5 - .83x_6 - .50x_7 + \varepsilon$$

where $x_1$ represents a male team before a female judge;

$x_2$, $x_4$ represent mixed teams before a male judge;

$x_3$,$x_5$ represent mixed teams before a female judge;

$x_6$ represents a female team before a male judge; and

$x_7$ represents a female team before a female judge.

If $y \geq .5$ is defined as an expected win, and $y < .5$ as an expected loss, for any given debate, then these results indicate that all-male teams had a greater expectation of winning before a male than before a female judge. Mixed teams and all-female teams, however, had an expected loss from male judges and an expected win from female judges.

When considering the dependent measure speaker rating, hypotheses 3.2 and 3.5 were rejected. The corrected form of Model #2 for prediction on the basis of sex was as follows.

$$\hat{y} = 22.80 - 3.3x_2 - 3.68x_4 - 3.47x_5 + \varepsilon,$$

This model indicates that the members of mixed teams received lower ratings than either all-male or all-female teams. Before a male judge, the expected speaker rating for the male member of a mixed team was 19.50, as compared to 22.80 for a male debater with a male colleague before a male judge. The expected rating for the female member of a mixed team before a male judge was 19.12. When debating before a female judge, the female in a mixed team had an expected rating of 19.33.

The dependent measure team rating also revealed significant effects due to sex. In general, female debaters tended to be associated with lower team ratings than did male debaters. Conversely, male judges tended to give lower team ratings than female judges.

The corrected form of Model #2 for predicting team rating on the basis of sex was as follows.

$$\hat{y} = 3.70 - .58x + \varepsilon,$$

when x represents mixed teams before a male judge.

The data also provided sufficient evidence to support the rejection of hypotheses 3.3 and 3.6 with $\alpha = .05$. The corrected form of Model #2 for predicting speaker rank by sex was as follows.

$$\hat{y} = 1.70 + .80x_1 + 1.43x_2 + 1.43x_4 +$$

$$.96x_5 + 1.30x_6 + \varepsilon$$

This model indicates that the members of mixed teams were ranked lower by male than by female judges. The expected rankings were 3.13 as compared to 1.70. For all-male teams, higher rankings came from male judges (1.7 as compared to 2.5 for female judges). For all-female teams, better rankings were received from female than from male judges.

The presence of significant interactions between the sex of the debaters and the sex of the judge represents a serious challenge to the integrity of intercollegiate debate. Immediate research is needed to discover means of compensating for the biases revealed

by this study. Until such means can be discovered, debate judges need to take rigorous steps to blot out whatever biases are affecting their decisions.

It should be noted in passing that this researcher did not accept the results of this study as indicative of a difference in performance between the sexes. Such a difference, however, if it did exist, might represent the cause of the significant regression effects discussed here. An experimental design which could hold ability constant while manipulating sex would shed more light on the issue, and would be of great value to the field of forensics.

The study also provided evidence of a significant effect due to the rater variable "prestige." While the overall measure, index of outcome, revealed no significant differences, each of the four measures from the A.F.A. Form "C" ballot indicated some effect due to prestige. Hypotheses 4.1, 4.2, and 4.3 were all rejected with $\alpha = .05$.

When considering the dependent measure win-loss, the corrected form of Model #3 for prediction on the basis of prestige was as follows.

$$\hat{y} = 1.00 - 1.00x_6 - .75x_7 - 1.00x_{11} - 1.00x_{12} - .83x_{15} \varepsilon,$$

where $x_6$ represents a low prestige team debating a high prestige team before a high prestige judge;

$x_7$ represents a low prestige team debating a high prestige team before a low prestige judge;

$x_{11}$ represents a low prestige team debating a junior college before a junior college judge;

$x_{12}$ represents a junior college team debating a high prestige team before a high prestige judge; and

$x_{15}$ represents a junior college team debating a low prestige team before a high prestige judge.

This model indicates that, except for the five cases mentioned above, there is no effect due to prestige in the win-loss outcome of intercollegiate debates. However, low prestige teams have an expected loss when debating high prestige teams before anything other than a junior college judge ($\hat{y} < .5$ is defined as an expected loss).

Low prestige teams also have an expected loss when debating a junior college team before a junior college judge. Junior college teams debating before a high prestige judge had an expected loss ($\hat{y} = 0.00, .17$) when debating anything other than another junior college team.

When considering the dependent measure speaker rating, the corrected form of Model #3 for prediction on the basis of prestige was as follows.

$$\hat{y} = 28.00 - 9.0x_1 - 7.17x_5 - 10.5x_7 - 8.63x_{10} - 10.5x_{11} - 6.83x_{13} - 7.83x_{14} - 10.0x_{15} - 8.13x_{16} - 9.13x_{17} + \varepsilon ,$$

where $x_1$ represents a high prestige team debating a low prestige team before a low prestige judge;

$x_5$ represents a high prestige team debating a junior college before a junior college judge;

$x_{10}$ represents a low prestige team debating a junior college before a low prestige judge;

$x_{13}$ represents a junior college team debating a high prestige team before a low prestige judge;

$x_{14}$ represents a junior college team debating a high prestige team before a junior college judge;

$x_{16}$ represents a junior college team debating a low prestige team before a low prestige judge; and

$x_{17}$ represents a junior college team debating a low prestige team before a junior college judge.

This model indicates that, for ten of eighteen possible combinations, prestige had an effect on speaker rating. High prestige teams debating low prestige teams before high prestige and junior college judges had an expected speaker rating of 28.0, as compared with an expected rating of 19.0 against a similar team but before

a low prestige judge. Against junior college teams, high prestige debaters had an expected rating of 28.0 from high and low prestige judges, but only 20.83 from junior college judges.

Against high prestige teams, low prestige debaters had an expected rating of 28.0 from high prestige and junior college judges, but only 17.5 from low prestige judges.

The rating scores of junior college debaters were the most affected by prestige. Against all low prestige teams, junior college debaters had an average expected rating of 18.92, with lowest scores coming from high prestige judges ($\hat{y}$ = 18.00) and highest scores coming from low prestige judges ($\hat{y}$ = 19.87). Against high prestige teams, junior college debaters received their highest scores from high prestige judges ($\hat{y}$ = 28.0), and their lowest scores from junior college judges ($\hat{y}$ = 21.17).

The effect of prestige on scores awarded by judges was most heavily concentrated in low prestige and junior college judges. In only one of six possible combinations did prestige have an effect on the scores given by high prestige judges ($x_{15}$). Low prestige judges were affected by prestige in five of the six possible combinations (omitting only $x_4$), and junior college judges showed an

effect in four of the six prestige combinations. In general, high prestige judges awarded higher speaker ratings than judges in either of the other two categories.

When considering the dependent measure team rating, prestige was found to have a significant effect in the rating for teams at all prestige levels as well as in the scores awarded by judges at all prestige levels. These are the data reflected in Table XVII in Chapter III. The corrected form of Model #3 for prediction of team ratings on the basis of prestige was as follows.

$$\hat{y} = 5.0 - 1.75x_1 - 1.00x_3 - 1.33x_4 - 1.67x_5 - 2.00x_7 - 1.33x_9 - 2.00x_{10} - 2.25x_{11} - 1.33x_{12} - 2.33x_{13} - 2.00x_{14} - 2.50x_{15} - 1.75x_{16} - 1.50x_{17} + \varepsilon,$$

where $x_3$ represents a high prestige team debating a junior college before a high prestige judge;

$x_4$ represents a high prestige team debating a junior college before a low prestige judge; and

$x_9$ represents a low prestige team debating a junior college before a high prestige judge.

This model indicates that in fourteen of eighteen possible combinations, prestige had a significant effect on team rating. High prestige teams had the highest

expected team ratings, with low prestige teams next. Junior college teams had the lowest expected ratings. The one combination where this did not hold involved ranking by a junior college judge.

Irrespective of the prestige level of the opposing team, high prestige teams received their highest ratings from high prestige judges, as did low prestige teams. Junior college teams, however, received their highest ratings from junior college judges.

Low prestige judges appeared to have reacted most to prestige. The ratings awarded by these judges showed significant regression in every one of the six combinations studied. High prestige judges and junior college judges were affected in four of the six combinations. High prestige judges awarded the highest team ratings in five of the six combinations. Junior college judges awarded the highest team ratings in three of the six combinations. (Two combinations received the same ratings from high prestige and junior college judges.) Low prestige judges tended to award the lowest team ratings.

As with the other dependent measures from the A.F.A. ballot, speaker rankings showed an effect due to prestige. The corrected form of Model #3 for prediction of speaker rank by prestige was as follows.

$$\hat{y} = 1.5 + 2.0x_6 + 1.75x_7 + 1.5x_8 + 1.38x_{11} +$$

$$1.67x_{12} + 1.75x_{15} + \varepsilon,$$

where $x_8$ represents a low prestige debater against a high prestige team before a junior college judge.

This model indicates that the rankings of high prestige debaters were unaffected by the prestige of the opposing team or the prestige of the judge. For all cases, the expected ranking of a high prestige debater was 1.50.

The rankings of low prestige debaters were affected by prestige in four of the six combinations studied. These debaters received their best rankings when debating junior college teams. The rankings of junior college debaters were affected by prestige in two of the six combinations studied. Significant differences appeared only when these debaters were speaking before a high prestige judge and, in those instances, junior college debaters received their worst rankings.

It must be noted that the results of this study cannot be taken as conclusive evidence that judges are affected by the prestige of the teams they hear debate. Differences in outcome may be caused by differences in the performance by the debaters. High prestige debaters may be better than low prestige or junior college debaters, and this may account for their superior results. However, the sample for this analysis was drawn only from elimination rounds and directly power-matched rounds.

In these cases, the performance levels of the two teams are theoretically very close. The presence of statistically significant differences is therefore disturbing. Further research is called for to increase awareness and understanding of the effects of prestige.

Because prestige is a rater variable, it may not be possible to eliminate its effects by manipulating tournament procedures. One practice which might have a limited success would be the elimination of school names from all tournament assignment sheets, and the use of randomly assigned team numbers to designate participants during a tournament. This is the practice at some tournaments, and it would be effective in those instances where the teams are not personally known by the judge in a round.

The results of this study indicate that the rater variable "proximity" also had some effect on the measures of the outcome of intercollegiate debates. In terms of the overall measure, index of outcome, the corrected form of Model #4 for prediction on the basis of proximity was as follows.

$$\hat{y} = -1.41 + 2.87x_1 + \varepsilon,$$

where $x_1$ represents the team farthest from the judge when both teams and the judge are from the same N.D.T. district.

This model indicated that, within any given N.D.T. district, proximity was a negative influence. Perhaps, in this case, proximity led to the growth of rivalries rather than friendships. Whatever the explanation, these results are counter to those from Brooks. For the dependent measure index of outcome, hypotheses 5.1 was rejected with $\alpha = .05$. However, there was insufficient evidence in the results to reject hypotheses 5.2, 5.3, or 5.4.

When considering the dependent measure win-loss, the findings were slightly different. No regression effects sufficient to reject any of the hypotheses were found. In the case of the dependent measure speaker rating, the corrected form of Model #4 for prediction on the basis of proximity was as follows.

$$\hat{y} = 22.25 - 4.65x_3 + \varepsilon,$$

where $x_3$ represents the team farthest from the judge when that judge comes from an N.D.T. district different from that of the two teams in the round.

This model supports the rejection of hypothesis 5.2 with $\alpha = .05$, and in this case offers support for the conclusions presented by Brooks.

The dependent measure speaker rank produced no significant regression effects. However, in the case of

team rating, there was an observable effect. The corrected form of Model #4 for predicting team rating on the basis of proximity was as follows.

$$\hat{y} = 3.83 - .63x_3 + \varepsilon,$$

This model contradicts the results of testing on speaker rating, and tends to oppose Brooks' conclusions.

The findings reported here offer conflicting evidence as to the effect of geographical proximity on the outcome of intercollegiate debates. In view of Brooks' more consistent results, additional research is needed to further explain the effects of this rater variable.

This study does provide consistent evidence that there is no effect due to N.D.T. district loyalties or animosities. In view of this, perhaps the National Debate Tournament is taking an unnecessary precaution by forbidding judges to hear a team from their own district against a team from another district.

For those effects of proximity which do exist, there seem to be two alternatives open to tournament directors. The first would involve balancing the geographical distance of the teams from the judge. The second procedure has already been discussed, i.e. using only team identification numbers in the course of a tournament.

It should be emphatically noted that the research reported here is only a beginning. Much more attention, especially experimentation, is needed in order to validate these results and to determine what other non-ability variables, if any, influence the outcome of intercollegiate debates. While essentially descriptive studies such as this one can be of great value in identifying problem areas, the capacity for testing solutions to those problems is uniquely a function of experimental research.

While the judges' perception of the sex of the debaters would be difficult to manipulate experimentally, both prestige and proximity would be, at least partially, susceptible of such manipulation. Experimentation which holds performance constant while varying the description of the source has a long history and extensive literature. Techniques should be available which would allow the manipulation of prestige and proximity as independent variables.

The competitive nature of forensics has, to this point in time, served to block any experimental treatments in situ. This must be overcome. The use of transcripts, video-taped debates, sophomore speech classes, etc., is of some value. But such vehicles are a poor substitute for the actual tournament situation.

Further study is needed of the component elements of the rater variable prestige. This study operationalized

that variable as a function of N.D.T. participation. Such a definition may be either inaccurate or incomplete. Only additional research can provide a comprehensive definition, and only experimental research can confirm the effects on the outcome of intercollegiate debates.

At present, debate tournaments rely primarily upon win-loss records and speaker ratings as measures of performance. The evidence of this research suggests that these measures are biased by the rater variables studied. Some new dependent measure is needed. The principal component analysis utilized here represented an attempt to develop such a measure. The results suggested, however, the possibility that factor scores tended to mask differences which did exist. Research is needed, first, to determine the proper bases for evaluation in debate and, second, to devise some overall measure which will serve as a valid and reliable index of outcome.

A number of statistical techniques are available which might aid in the search for such a measure. Factor analysis should not be abandoned on the basis of this study. It merits further and intensive examination. Various correlation techniques, if studied, might lead to the development of some index expressed in terms of its relationship to current measures.

This study presents evidence of the validity of the debate format currently in use. The evidences of rater bias which did emerge are most likely beyond the control of the average tournament director. There are, however, some possible solutions suggested by the results of this study.

Given access to high-speed computing facilities, individual and team scores on the present ballot could be adjusted on the basis of the statistical models developed here. This would compensate for the effects of rater bias. In effect, it would involve the creation of a handicapping procedure for debate. The number of possible models, and their complexity, would demand automatic computation to avoid unreasonable delays. While this solution seems logically based, it would probably be received with great resistance and hostility on the debate circuit. Coaches and debaters would probably resent, and be suspicious of, any "tinkering" with their scores. Acceptance of such a solution would depend upon absolute proof that rater bias was affecting outcome. Such proof can only come through experimental research. Thus the need for more experimentation!

Given sufficient funds, "outside" judges could be hired. It would be possible to bring in debate coaches from different areas of the country who do not have

teams participating in the tournament. This has been suggested for the various N.D.T. district tournaments, but no substantive action has been taken. Expense remains a major barrier. Outside, non-debate, judges are an additional possibility. Local business and professional persons, faculty from the host school, etc., could be recruited to serve as judges. It cannot be assumed, however, that rater bias would not affect these judges just as it appears to affect coaches who now serve as judges. Research would be needed to determine whether or not rater bias would operate among such a group.

The ultimate solution to the problems of rater bias must be found in the judges who render the decisions. Better training for debate coaches, emphasizing the proper basis for awarding decisions, would be of value. This is an area in which the proposed Developmental Conference on Forensics should act. Such a conference could and should propose specific standards of competence for debate coaches. It could also initiate and support financially the research needed to determine and empirically verify a set of criteria for evaluating debate.

Some corrective influence would probably be generated by publicity of the results of this and

similar studies. Rater bias will not go away simply because it is ignored. The opening assembly of every debate tournament should include a discussion of rater variables. Biases should be identified specifically, and judges should be encouraged to be sensitive to the influence of such biases.

A stronger, more precise statement of professional ethics might be useful in conjunction with this practice. Some procedure for investigating and eliminating consistently biased judges might have beneficial results. All of these are areas in which the American Forensics Association can and should exercise its leadership.

BIBLIOGRAPHY

## Books


Dixon, W.J. BMD Biomedical Computer Programs. University of California Publications in Automatic Computation No. 3, 1972.

_____. BMD Biomedical Computer Programs: X-series Supplement. University of California Publications in Automatic Computation No. 3, 1972.

Hayes, William. Statistics. New York: Holt, Rinehart and Winston, 1968.

Mendenhall, William, Lyman Ott, and Richard Schaeffer. Elementary Survey Sampling. Belmont, California: Wadsworth Publishing Company, Inc., 1971.

Mills, Glen E. Argumentation and Debate: Techniques of a Free Society. New York: The Macmillan Company, 1964.


## Journals


Allen, R.R. "The Effects of Interpersonal and Concept Compatibility on the Encoding Behavior and Achievement of Debate Teams." Central States Speech Journal, XIV (February, 1963), 23-26.

Benson, James A. "The Use of Evidence in Intercollegiate Debate." Journal of the American Forensics Association, VII (Spring, 1971), 260-270.

Brooks, William D. "Judging Bias In Intercollegiate Debate." Journal of the American Forensics Association, VII (Winter, 1971), 197-200.

Cathcart, Robert S. "An Experimental Study of the Relative Effectiveness of Four Methods of Presenting Evidence." Speech Monographs, XXII (August, 1955), 227-233.

Conklin, Forrest (ed.) "A Bibliography of Argumentation and Debate for 1969." Journal of the American Forensics Association, VIII (Fall, 1971), 81-104.

Cragen, John F. and Donald C. Shields. "The Comparative Advantage Negative." Journal of the American Forensics Association, VII (Spring, 1970), 85-91.

Douglas, Donald G. "Forensics Studies in Contemporary Speech Education." Journal of the American Forensics Association, VIII (Spring, 1972), 178-181.

_____. "Toward a Philosophy of Forensics Education." Journal of the American Forensics Association, VIII (Summer, 1971), 36-41.

Dresser, William R. "The Use of Evidence in Ten Championship Debates." Journal of the American Forensics Association, I (September, 1964), 101-106.

Ellis, Dean S. and Robert Minter. "How Good Are Debate Judges?" Journal of the American Forensics Association, IV (Spring, 1967), 53-56.

Faules, Don. "Measuring Refutation Skills: An Exploratory Study." Journal of the American Forensics Association, IV (Spring, 1967), 47-52.

Granell, Lee. "Forensics and the Department of Speech Communication." Journal of the American Forensics Association, VIII (Spring, 1972), 186-188.

Hayes, Michael T. and Joe McAdoo. "Debate Performance: Differences Between Male and Female Rankings." Journal of the American Forensics Association, VIII (Winter, 1972), 127-131.

Hensley, Wayne E. "A Profile of the N.F.L. High School Forensics Director." Journal of the American Forensics Association, IX (Summer, 1972), 282-287.

Hensley, Wayne E. and David B. Strother. "Success in Debate." The Speech Teacher, XVII (September, 1968), 235-237.

Hill, Sidney R., Jr. "A Study of Participant Evaluations in Debate." Journal of the American Forensics Association, IX (Winter, 1973), 371-377.

Hufford, Roger, "Toward Improved Tournament Judging." Journal of the American Forensics Association, II (September, 1965), 120-125.

Kennedy, Allan J. "Directory of Universities and Colleges Conducting Summer Institutes." Journal of the American Forensics Association, VII (Winter, 1971), 224-233.

Kruger, Arthur N. "The Comparative Advantage Case: A Disadvantage." Journal of the American Forensics Association, III (September, 1966), 204-211.

_____. "Judging the Judging at Meadville." Bulletin of the Debating Association of Pennsylvania Colleges, XII (December, 1955), 32.

Kully, Robert D. "Forensics and the Speech Communication Discipline." Journal of the American Forensics Association, VIII (Spring, 1972), 192-199.

Larson, Carl E. and Kim Giffin. "Ethical Consideration in the Attitudes and Practices of College Debaters." Journal of the American Forensics Association, I (September, 1964), 86-90.

Ling, David A. and Robert V. Seltzer. "The Role of Attitudinal Inherency in Contempory Debate." Journal of the American Forensics Association, VII (Spring, 1971), 271-277.

McGuckin, Henry. "Better Forensics: An Impossible Dream?" Journal of the American Forensics Association, VIII (Spring, 1972), 182-185.

Marcham, Frederick George. "Teaching Critical Thinking and the Use of Evidence." Quarterly Journal of Speech, XXXI (October, 1945), 362-368.

Melzer, Dorothy. "Suggestions for Improving Debate Judging." Southern Speech Journal, XVIII (1952), 51-63.

"Minutes of the 24th Annual American Forensics Association Meeting." Journal of the American Forensics Association, IX (Winter, 1973), 380-391.

Newman, Robert P. and Keith R. Sanders. "A Study in the Integrity of Evidence." Journal of the American Forensics Association, II (January, 1965), 7-13.

Phillips, Gerald M. "Experimentation and the Future of Debate." The Gavel, XLV (November, 1962), 3-6.

Shiffrin, Steven. "Forensics, Dialectic and Speech Communication." Journal of the American Forensics Association, VIII (Spring, 1972), 189-191.

Shrum, Robert. "Do Judges Know When Debaters Win or Lose?" Rostrum, XLII (1968), 6-7.

Smith, Harold E. "The Use of Statistical Evidence in Debate." Quarterly Journal of Speech, XXVI(October, 1940), 426-431.

Wenzel, Joseph W. "Campus and Community Programs in Forensics." Journal of the American Forensics Association, VII (Spring, 1971), 253-259.

Wilmington, S. Clay. "A Study of the Relationships of Selected Factors to Debate Effectiveness." Central States Speech Journal, XX(Spring, 1969), 36-39.

Wise, Charles N. "Relationships Among Certain Rating Judgments on the Form C. Ballot." Journal of the American Forensics Association, VIII (Spring, 1971), 305-308.

## Theses and Dissertations

Brembeck, Winston L. "The Effects of a Course in Argumentation on Critical Thinking Ability." Ph.D. dissertation, University of Wisconsin, 1947.

Cromwell, Harvey. "The Relative Effect on Audience Attitude of the First Versus the Second Argumentative Speech of a Series." Ph.D. dissertation, Purdue University, 1949.

Dollar, David L. "An Examination of the Importance Assigned Criteria of Debate Evaluations." M.S. thesis, Kansas State Teachers College at Emporia, 1968.

Howell, William Smiley. "The Effects of High School Debating on Critical Thinking." Ph.D. dissertation, University of Wisconsin, 1942.

Keeling, Russell M. "An Analysis of Refutation and Rebuttal in Interscholastic Debate." M.A. thesis, Baylor University, 1959.

Madsen, Sandra. "A Study of the Use of Specified Criteria . . . in Evaluating Tournament Debates." M.A. thesis, Wisconsin State University, 1969.

Mills, Beatrice Bahr. "A Survey of Graduate Research in Debate." M.A. thesis, University of Southern California, 1952.

## BIOGRAPHICAL SKETCH

Sidney Ray Hill, Jr., was born on 9 December 1943 in Biloxi, Mississippi. He is the only son of Sidney Ray and Sue Norton Hill, both native Alabamians. He was educated in the public schools of Birmingham and Jefferson County, Alabama.

In September of 1961, Sidney Hill matriculated into the pre-law curriculum of the University of Alabama. In 1964 he transferred to Birmingham-Southern College, from which he was graduated in 1965 with a major in history. His graduate education was received in the Departments of Speech of Auburn University and the University of Florida.

From September 1965, until June, 1966, Sidney Hill was employed as a graduate assistant in the Department of Speech at Auburn University. During the academic year 1966-67, he held an appointment as Instructor and Director of Forensics at Birmingham-Southern College. At the end of this appointment, Mr. Hill returned to Auburn University where he served as Director of Forensics. Since September of 1969, he has served as a Graduate Teaching Assistant in the Department of Speech at the University of Florida, while completing requirements for his Ph.D. degree.

Sidney Hill is married to the former Margaret Ruth Bailey of Knoxville, Tennessee. They have no children. He is a member of various national and regional speech associations, Delta Sigma Rho-Tau Kappa Alpha honor fraternity, and Chi Phi social fraternity.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Douglas G. Bock, Chairman
Assistant Professor of Speech

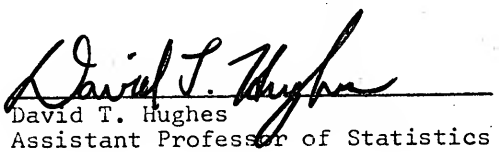
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Donald E. Williams
Professor of Speech

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Anthony J. Clark
Assistant Professor of Speech

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



David T. Hughes
Assistant Professor of Statistics

This dissertation was submitted to the Department of Speech in the College of Arts and Sciences and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dean, Graduate School

August, 1973